Journal of Pre-College Engineering Education Research (J-PEER)

Volume 10 | Issue 2 Article 4

2020

Examining the Factor Structure of a Middle School STEM Occupational Values Scale

Jessica Harlan
Gallup, Inc., jessica.harlan@jhmi.edu

James Van Haneghan
University of South Alabama, jvanhane@southalabama.edu

Recommended Citation

Harlan, J., & Van Haneghan, J. (2020). Examining the Factor Structure of a Middle School STEM Occupational Values Scale. *Journal of Pre-College Engineering Education Research (J-PEER), 10*(2), Article 4.
https://doi.org/10.7771/2157-9288.1141

Examining the Factor Structure of a Middle School STEM Occupational Values Scale

Abstract

As part of a longitudinal study of the development and implementation of a middle school engineering design curriculum, we have used an occupational values subscale of the Assessing Men and Women in Engineering (AWE) project's Engineering version of the Core Survey for Middle School-Aged Participants to measure student occupational interest in science, technology, engineering, and mathematics (STEM). According to the developers, this set of tools is intended to measure factors related to STEM careers, including occupational interests. While the AWE tools have been widely used, there have been no formal examinations of the psychometric properties of the middle school tools. Using a sample of our program participants, we examined the underlying factor structures of the occupational values subscale using exploratory and confirmatory factor analysis. We found that the AWE Work Values scale assesses two separate sets of occupational values: (1) using analytical and problem-solving skills and (2) personal satisfaction. Even though these two factors were confirmed, we conclude that there is still a need to improve reliability and more clearly define the constructs measured.

Keywords

occupational values, psychometric, validity, middle grades, integrated STEM, program evaluation

Document Type

Article

Cover Page Footnote

This report was partially supported by NSF Award DRL- 0918769. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Available online at http://docs.lib.purdue.edu/jpeer

Journal of Pre-College Engineering Education Research 10:2 (2020) 47–57

Examining the Factor Structure of a Middle School STEM Occupational Values Scale

Jessica Harlan[1] and James Van Haneghan[2]

[1]*Gallup, Inc.*
[2]*University of South Alabama*

Abstract

As part of a longitudinal study of the development and implementation of a middle school engineering design curriculum, we have used an occupational values subscale of the Assessing Men and Women in Engineering (AWE) project's Engineering version of the Core Survey for Middle School-Aged Participants to measure student occupational interest in science, technology, engineering, and mathematics (STEM). According to the developers, this set of tools is intended to measure factors related to STEM careers, including occupational interests. While the AWE tools have been widely used, there have been no formal examinations of the psychometric properties of the middle school tools. Using a sample of our program participants, we examined the underlying factor structures of the occupational values subscale using exploratory and confirmatory factor analysis. We found that the AWE Work Values scale assesses two separate sets of occupational values: (1) using analytical and problem-solving skills and (2) personal satisfaction. Even though these two factors were confirmed, we conclude that there is still a need to improve reliability and more clearly define the constructs measured.

Keywords: occupational values, psychometric, validity, middle grades, integrated STEM, program evaluation

Introduction

There has been and continues to be significant discussion in the United States about the need to increase the pipeline of highly qualified workers in science, technology, engineering, and mathematics (STEM) fields (Deloitte Consulting LLP et al., 2009; Maltese & Tai, 2011). To address this need, the federal government, foundations, and industry have invested millions of dollars to increase workforce readiness in STEM among K–12 students in the hope that these students will choose to enter and remain in the STEM workforce (Gonzalez, 2012; STEM Education, 2013). Many of these programs used an approach generally referred to as "integrated STEM," which involves presenting concepts from one STEM area within the context of another (Honey et al., 2014). For example, engineering principles might be presented and practiced in the context of science education. This approach is often intended to trigger or build an interest in both STEM content and STEM careers.

As the field of integrated STEM education has grown, so has the need for assessments. However, despite the growth of integrated STEM programs, the development and improvement of assessment tools have not kept pace (Honey et al., 2014). This lack of strong measurement tools has not eliminated the need for STEM education programs to measure program outcomes. However, given the significant time and resources required to develop a new, psychometrically tested occupational values scale related specifically to STEM occupations, program evaluators often elect to use existing and often insufficient tools rather than develop new ones. This was the case for the project we were evaluating.

In the 2008–2009 school year, collaborative work began with a local education foundation, university, and school district to design, develop, and implement a middle school integrated STEM program. This program used mathematics and science classroom-based engineering design experiences to help middle school students gain greater STEM knowledge and skill (Harlan et al., 2014). In 2008–2009, the concept of integrated STEM was new and uncommon in pre-high school environments. Then, as now, there were few tools to measure integrated STEM program outcomes. On our project, the short time between program funding and program implementation made it impractical to develop and test a new tool. To assess outcomes related to student attitudes and beliefs (including occupational values), the original evaluator for this project chose to use the Assessing Women and Men in Engineering (AWE) project's AWE Engineering version of the Core Survey for Middle School-Aged Participants (AWE Middle School Survey) (AWE, 2008).

Assessing Women and Men in Engineering

The AWE tools, including the Pre-College Surveys, were developed as part of an NSF-funded, multi-institutional collaboration to develop and test assessment and evaluation tools related to STEM education (AWE, 2008). According to the AWE developers, the Pre-College Surveys are intended to measure factors related to STEM careers, including interests, skill and confidence, knowledge, and career plans. Although the developers do not specify which subscales of the Pre-College Surveys are intended to map which constructs, they state that the middle and high school versions of the tool measure:

> Course-taking plans for high school, whether participant intends to study science, engineering, or computer [sic], what participant knows about what engineers, scientists, or computer scientists do, what factors (if any) about being an engineer, scientist, or computer scientist appeal to participant, events or persons that influenced participants' study plans, participant skill and confidence level in areas that are important for successfully completing a science, engineering, or computer degree, where participants plan to study science, engineering, or computer science/engineering, her/his satisfaction with the quality of the activity in which she/he has participated. (AWE, 2008)

Items on the AWE Middle School Survey presumably intended to measure occupational values for STEM fields (AWE Work Values scale) seem to align with some common constructs, such as family and power. However, the AWE Work Values scale also has additional items related primarily to STEM careers, such as "Work that allows me to use math, computer, engineering, or science skills." This seems to indicate that the scale might provide insight into student interest in STEM careers specifically.

The AWE Pre-College Surveys have been widely used over the last several years (e.g., Brevik et al., 2015; Demetry & Sontgerath, 2013). Tools such as the AWE Middle School and other surveys, which are available online at no cost, are often appealing to organizations that do not have the resources to build their own tools or pay for a psychometrically tested tool. The tools are also recommended by organizations that support the development of STEM programs and interventions (e.g., National Girls Collaborative Project, n.d.; Rochester Institute of Technology, n.d.).

Despite their popularity, there is little literature examining the psychometric properties of the AWE Middle School tools. The developers tested the Pre-College Surveys with experts and audiences, but no formal validation studies were performed on these tools by the developers (AWE, n.d.a). After extensive searches, the only article we were able to find was one presented by one of the authors at a research summit that used an exploratory factor analysis to examine the factor structure of the AWE Work Values scale using initial program data (Van Haneghan et al., 2012).

Purpose of the Research

Despite the significant body of research on how to develop attitude scales, there is little information available in the literature about measurement tools that are currently in use by those researching and evaluating STEM programs. It is important to address the psychometric properties of these tools to aid their users in understanding the tools' qualities and limitations. If researchers find strong evidence that these tools are reliable and valid, users can feel more confident when evaluating program outcomes. If researchers find that these tools have problems, however, such information can also aid users in identifying limitations for interpretation and data use (for example, when making high-stakes program decisions). Additionally, such investigation can serve as a foundation for researchers, evaluators, and other practitioners to improve the quality of tools available. Knowing what types of individual items and groupings of items seem to measure a particular construct can provide those developing similar tools with information to limit the extent to which they must begin from scratch.

The present research seeks to examine the factor structure of the AWE Work Values subscale of the AWE Middle School Survey. Using data from students who responded to the AWE Work Values scale over multiple years, we examined the extent to which the factor structure showed evidence of measuring occupational values related to STEM careers.

Methods

Participants

Participants were 1,754 students who were part of a longitudinal, quasi-experimental study of the development and implementation of a middle school engineering design curriculum. We collected data using the AWE Middle School Survey in Grade 6 and again in Grade 8. Data from the larger study were collected from the 2008–2009 school year until the 2013–2014 school year. The present research uses Grade 8 data from students who received the final version of the scale under investigation. These students completed Grade 8 in the 2011–2012 (n = 648), 2012–2013 (n = 565), and 2013–2014 (n = 541) school years. For simplicity, we will refer to cohorts of students by the year in which they would have completed Grade 8.

Using SPSS (Version 22), we obtained four randomly selected samples from the total pool of students. To maximize the power of our confirmatory factor analyses, we elected to use a smaller sample size for our exploratory factor analysis. Sample 1 (n = 173) was used to perform the exploratory factor analysis. This sample size was consistent with recommendations that suggest that, with a small number of factors and large communalities, only 100 or so cases are needed (Bandolos & Finny, 2010). This sample represented approximately 10% of the total population of data. The remaining 1,581 students were randomly divided into about thirds, with 553 students in Sample 2, 540 in Sample 3, and 488 in Sample 4. Sample 2 was used to test Model A, Sample 3 was used to test Model B, and Sample 4 was used to test whether the findings for Model B could be replicated.

Data were collected at a math and science magnet school and two regular middle schools. The magnet school (TMag) and one of the regular middle schools (TReg) were implementing the engineering curriculum, while the other middle school was a comparison school (CReg) that did not implement the curriculum. The proportion of students from each type of school did not differ significantly by sample (X^2 (6, N = 1,680) = 4.76, p = 0.58). About 5–9% of each sample were students from TMag. About 45–49% of each sample were students from TReg. Lastly, about 43–48% of each sample were students from CReg.

We examined several variables to ensure that, while the schools were not identical in focus, they did not systematically differ in make-up. All schools were Title I schools, and over half of the students at each school received free lunch. All three schools had similar proportions of male and female students (Table 1). Gender and race/ethnicity did not vary significantly across years.

We also examined the standardized achievement scores of students at each of the schools, using each individual sample and the overall sample where they were available. Scores on this criterion-referenced state test can range from 200 to 800. In Grade 7, for example, the student would be classified as meeting or exceeding proficiency with a reading score of 640 or higher and a mathematics score of 651 or higher.

Reading and mathematics achievement scores from all schools and samples were available for Grade 5. In Grades 6, 7, and 8, only reading and mathematics scores were available for TReg and CReg schools. We analyzed the Grade 5 scores for students from all three schools using a 3 (school) × 2 (sample) ANOVA to examine whether their scores varied by four

Table 1
Demographic characteristics of the sample by school and by sample group.

	Achievement scores		Gender		Race/ethnicity				
	Reading	Math	Female	Male	Asian/Pacific American	Black/African American	Latina/o/Hispanic American	White	Other
Overall	**675.93**	**668.26**	**49%**	**51%**	**3%**	**41%**	**3%**	**49%**	**4%**
TMag (n = 108)	691.00	713.00	61%	39%	5%	44%	2%	45%	5%
TReg (n = 827)	675.21	672.27	48%	52%	2%	52%	3%	40%	4%
CReg (n = 819)	676.50	663.59	48%	52%	4%	29%	4%	58%	5%
Sample 1 (n = 173)	671.23	664.93	57%	43%	1%	36%	2%	55%	5%
Sample 2 (n = 531)	674.92	667.04	48%	52%	4%	42%	3%	48%	4%
Sample 3 (n = 511)	678.80	672.91	48%	52%	4%	44%	4%	44%	4%
Sample 4 (n = 465)	676.97	667.90	49%	51%	3%	36%	3%	54%	5%
Missing data (removed, n = 74)	668.55	656.80	43%	57%	3%	52%	1%	39%	5%

samples (one exploratory sample and three model testing samples). We examined the differences between samples using a 2 (school) × 4 (sample) ANOVA for the data from the TReg and CReg schools for Grades 6, 7, and 8, as we did not have the data for the TMag school. Data were available for most students in the samples. Given the random draw of students into samples, we did not anticipate there would be any interaction between school and sample that could create issues in replicating results.

As we expected, we found no differences between the samples, nor any interaction between sample and school. We did find that the magnet school students (TMag) had higher achievement scores compared to the regular school students (TReg and CReg). This was consistent with what we know about the schools' overall performances. Most importantly, we did not find any differences in the achievement levels of the four samples at Grade 5. In looking at data from Grades 6–8, which only consisted of scores in the four samples from the TReg and CReg schools, we found that there were no consistent main effects of nor, more importantly, any interactions related to achievement. We felt confident that the samples were equivalent and were comfortable generalizing results across them. We do not include details of these analyses here, but the results of the analyses will be made available to interested readers.

We used IBM SPSS AMOS Version 22 to perform our confirmatory factor analyses. We decided that since we had small amounts of missing data, we would use listwise deletion. We included students from Samples 2, 3, and 4 (those used for confirmatory factor analyses) who had provided a response to each item on the scale. There were 4% of students in Sample 2 with missing data, 5% from Sample 3, and 5% from Sample 4. As noted above, students with missing data had slightly lower achievement scores than students without missing data (Table 1). Additionally, a higher proportion of students with missing data were male, and a higher proportion identified their race/ethnicity as Black. We performed a logistic regression and found that none of these factors were significant predictors of whether a student would have missing data (and thus be excluded), X^2 (4, $N = 1{,}581$) = 6.62, $p = 0.158$. Given this information, we elected not to use any data substitution methods for missing data because of the low proportion of students with missing data from each sample. Rather, those with missing data were excluded from analyses.

Instrument

The Assessing Women and Men in Engineering Project was an NSF-funded group, primarily from the Pennsylvania State University and University of Missouri (AWE, n.d.b). It was developed in 2001 and continued through 2009. This group developed quantitative assessment tools for use in research and evaluations of K–16 formal and informal education and outreach programs. They subsequently made these tools available for free from the group's website.

As mentioned previously, the present research examines the factor structure of a subscale of the AWE Middle School Survey that the developers seem to have intended to measure occupational values for STEM fields (Work Values scale). These items appear on each of the three time-point versions (Pre-Participation, Immediate Post-Participation, and 3–6 Month Post-Participation) for each content area version of the survey (Engineering, Science, Computer Science). This 10-item scale originally used a 3-point, fully anchored scale (i.e., each numeric scale point had an adjective that indicated the importance specified by a rating—not important, somewhat important, very important). However, in our early use of the tool, we noted that there was little variability in responses using the 3-point scale. The data from students who responded using the 3-point scale were the data used in Van Haneghan et al.'s (2012) initial examination of the scale's factor structure. When we updated the questionnaire before collecting Grade 8 data in the 2011–2012 school year, we modified the scaling for the Work Values items to be a 4-point, fully anchored scale (not important, somewhat important, important, very important). The students included in the present research responded to these items on a 4-point scale rather than the original 3-point scale.

There was one item on the scale ("Work that is fun") that we felt was not a good discriminator for children at this age. Over 80% of students responded to this item with "important" or "very important." For this reason, we determined that the item did not contribute additional information to the model and did not include it in our analyses.

Results and Discussion

Exploratory Factor Analysis

Using exploratory factor analysis, we examined the factor structure of the Work Values scale using SPSS Version 22 in two ways. We first used a principal axis extraction with a Varimax rotation to examine the factor structure across subgroups. We only examined factors with eigenvalues of 1 or greater, and suppressed loadings lower than 0.35. This cutoff fitted with the suggested rule of thumb that a loading account for at least 10% of the variance of a variable (Pituch & Stevens, 2016). As can be seen in Table 2, the items loaded on two factors. As can also be seen in Table 2, there is one item that cross loaded on two factors, with little difference between the factor loadings. We eliminated this item and re-ran the analysis.

Table 2
Exploratory factor analysis (EFA) of the Work Values scale: factor loadings across subgroups.

	Original EFA		Modified EFA	
	Factor 1	**Factor 2**	**Factor 1**	**Factor 2**
Work that makes me think		0.61		0.59
Work that allows me to make lots of money	0.72		0.72	
Work that allows me to use math, computer, engineering, or science skills		0.74		0.75
Work that allows me to tell other people what to do	0.36		0.36	
Work that allows me to help solve problems and create solutions		0.58		0.58
Work that allows me to have time with family	0.51		0.44	
Work that allows me to help my community and/or society	0.44	0.36	—	—
Work that makes people think highly of me	0.56		0.64	
Work that is satisfying to me	0.62		0.55	

Table 3
Parallel analysis raw data eigenvalues and 95th percentile random data eigenvalues.

Root	Raw data eigenvalues	95th percentile eigenvalues
1	1.879	0.519
2	0.794	0.352
3	0.079	0.234

The items on Factor 1 appeared to be measuring the importance of personal satisfaction. These values could be present in any field, including but not limited to STEM careers. They are similar to the items included in many other work values scales (e.g., Eccles, 1994; Weisgram & Bigler, 2006). Factor 2 appeared to be measuring the importance of using analytical and problem-solving skills. The valuing of items on this factor would be consistent with work activities in many fields, including jobs that comprise the STEM field. This is discussed further in the Conclusions section of this paper.

We next examined the factor structure using parallel analysis (O'Connor, 2000). Using the rawpar syntax from O'Connor (2000), we ran a principal axis parallel analysis using 1,000 permutations of the original raw data set. According to O'Connor, the permutations method, which uses Castellan's (1992) algorithm, is more robust than using the normally distributed random data generation parallel analysis method. We compared the parallel analysis eigenvalues for the 95th percentile to those of the raw data for the 8-variable Work Values scale with 166 valid cases. A comparison of the raw data eigenvalues to the 95th percentile eigenvalues supported our initial finding that this scale is, in fact, measuring two factors (Table 3).

While the exploratory two-factor model appears to fit the data, the clarity of the model is muddled by the need to remove one of the items that loaded on both factors. The item concerning "Work that helps the community or society" is a potentially important value for individuals in STEM fields. Additionally, Factor 1, the personal satisfaction factor, pulls together a number of different motivations and values. Even though it holds together statistically, its clarity as a unifying set of values is not optimal. Having at least shown that two factors can be generated from this set of items, we move on to examine them from a confirmatory perspective using IBM SPSS AMOS Version 22.

Confirmatory Factor Analysis: Two-Factor Model

We performed multiple confirmatory factor analyses using IBM SPSS AMOS Version 22. We hypothesized a two-factor model (Figure 1) to be confirmed in Grade 8. We found a small positive correlation ($r = 0.234$, $p = 0.002$) between the two factors in our initial model development. For this reason, Model A included a covariance between the two factors. To determine model fit, we compared goodness-of-fit statistics for each tested model and the replication to recommended values, including comparative fit index (CFI) > 0.95 (Hu & Bentler, 1999; Schreiber et al., 2006), root mean square of approximation (RMSEA) < 0.05 (Byrne, 2010), and standardized root mean square residual (SRMR) < 0.08 (Hu & Bentler, 1999).

For Model A, we specified 9 regressions, 9 variances, and 1 covariance totaling 19 parameters. For the confirmatory factor analysis of Model A, we had a sample size of 553, for a ratio of 29.10 participants to each 1 parameter estimated. For our final model, Model B, we specified 9 regressions, 9 variances, and 2 covariances, totaling 20 parameters. With a sample size of 540, we had a ratio of participants to estimated parameters of 27 to 1. In our replication of Model B, we had a sample size of 488, for a ratio of participants to estimated parameters of 24.4.

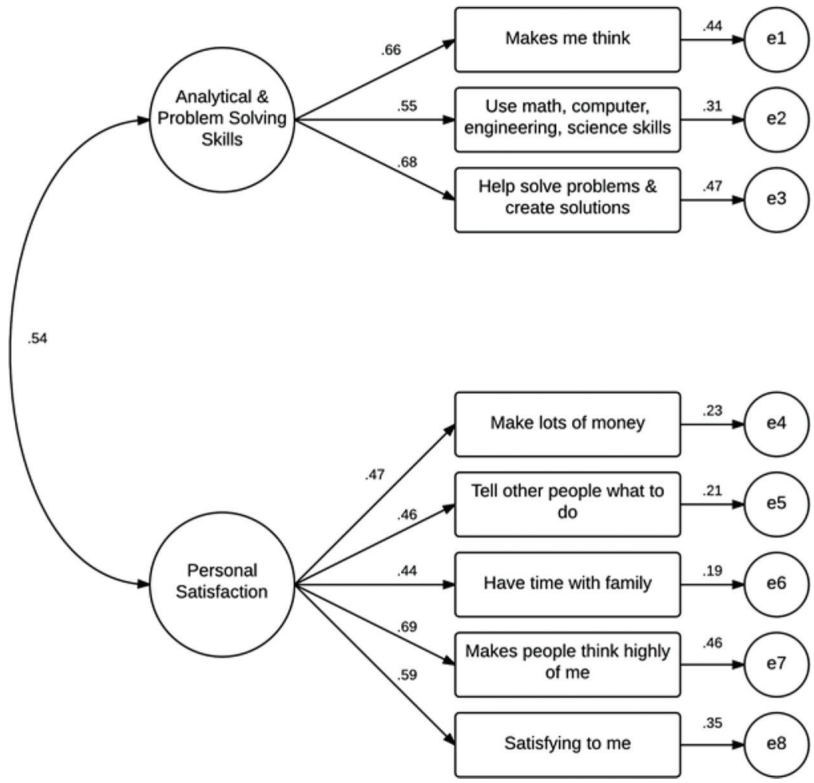

Figure 1. Path coefficients for Model A (the hypothesized model).

Table 4
Summary goodness-of-fit statistics in determination of baseline models.

Model	X^2	*df*	CFI	RMSEA	SRMR
Model A (hypothesized model)	65.835	19	0.930	0.068	0.0436
Model B with one error covariance (e5 and e7)	37.703	18	0.968	0.046	0.0329
Model B replication	19.737	18	0.995	0.014	0.0327

Using an examination of the z-scores, we detected outliers on the item "Work that allows me to make lots of money" in Samples 2, 3, and 4. In Sample 2, $z = -3.65$ for 2.3% of cases on this item. In Sample 3, $z = -3.80$ for 2.3% of cases on this item. In Sample 4, $z = -4.08$ in 1.3% of cases. In Samples 3 and 4, we also detected outliers on the items "Work that allows me to have time with my family" ($z = -2.72$ in 4.3% of cases in Sample 3; $z = -2.54$ in 4.7% of cases in Sample 4) and "Work that is satisfying to me" ($z = -3.47$ in 2.3% of cases in Sample 3; $z = -3.11$ in 3.2% of cases in Sample 4). Given the small proportion of cases with outliers, we elected to keep them in the sample.

The goodness-of-fit indices for the hypothesized model, Model A, did not meet the cutoff criteria established for this study (Table 4).

To improve the model, we examined modification indices. For the hypothesized model, the covariance modification index for Model A was 11.40 for the error terms associated with the items "Work that allows me to tell other people what to do" and "Work that makes people think highly of me." Given that these items both seem to be measuring some level of desire for power and prestige, we would expect there to be some level of covariance between these two variables. To improve model fit, we developed Model B by adding a covariance to the error terms e5 and e7.

Model B (Figure 2) showed good model fit consistent with the parameters recommended by Hu and Bentler (1999), Schreiber et al. (2006), and Byrne (2010).

Using Sample 4, we tested to see whether we could replicate good model fit for Model B with a new sample (Figure 3). Once again, Model B showed good model fit using the criteria established by Hu and Bentler (1999), Schreiber et al. (2006), and Byrne (2010).

See Table 4 for goodness-of-fit indices and Table 5 for factor loadings for both Model B analyses.

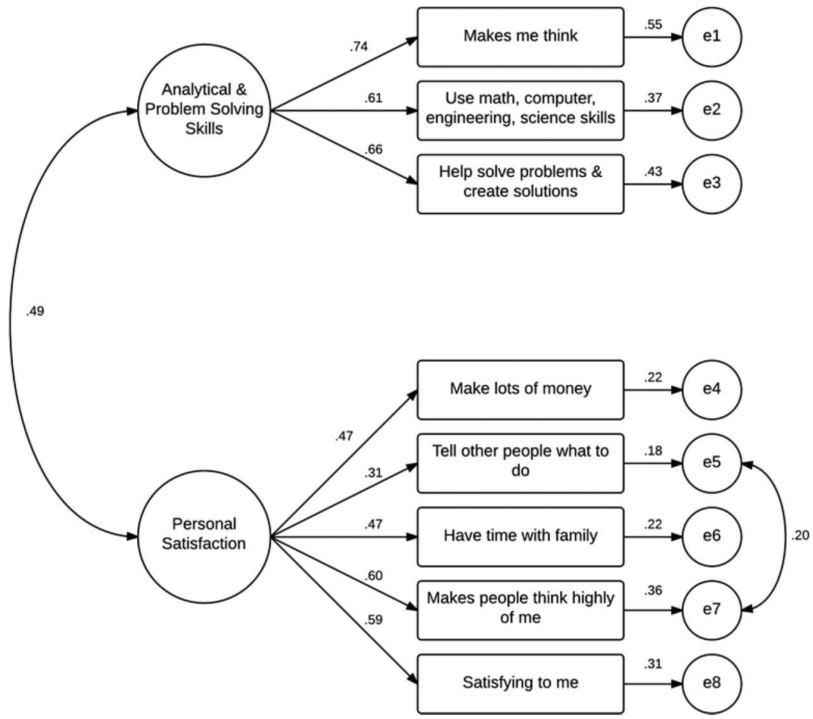

Figure 2. Path coefficients for Model B, which includes one additional covariance.

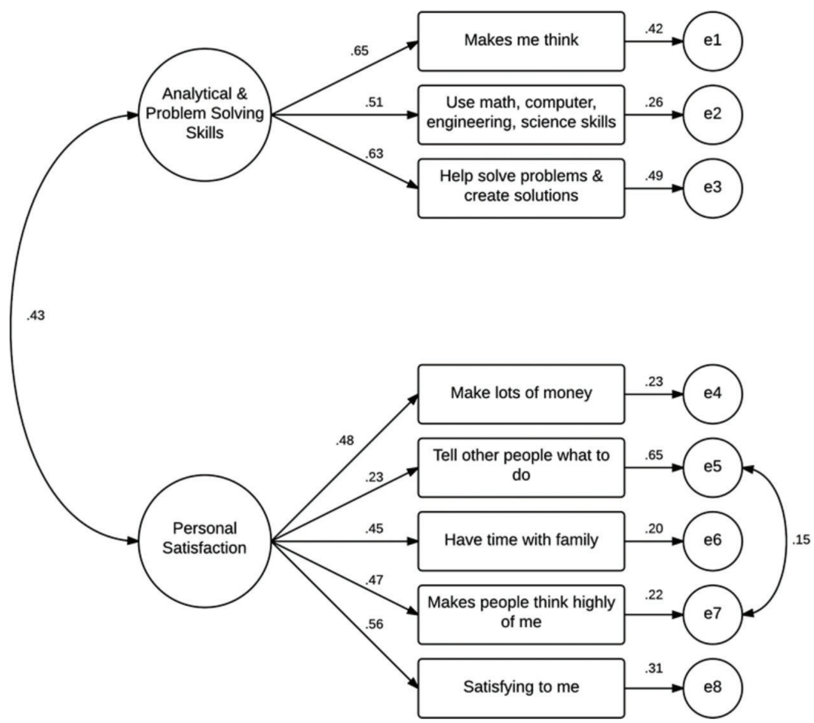

Figure 3. Path coefficients for the replication of Model B (the final model).

Scale Reliability

Following our examination of factor structure, we examined the internal consistency reliability of the items as originally presented by the developers, as well as each factor we identified across samples. The internal consistency reliability of all items together was $\alpha = 0.74$. This is lower than the generally accepted minimum standard of $\alpha = 0.80$. The internal consistency reliability of the "Analysis and Problem Solving" scale was $\alpha = 0.66$. The "Personal Satisfaction" scale had an internal

Table 5
Standardized and unstandardized coefficients for Model B.

		Original analysis			Replication		
Observed variable	Latent construct	*β*	*B*	SE	*β*	*B*	SE
Work that makes me think	Analysis and Problem Solving	0.745	1.00		0.650	1.00	
Work that allows me to use math, computer, engineering, or science skills	Analysis and Problem Solving	0.609	0.97	0.10	0.511	0.94	0.13
Work that allows me to help solve problems and create solutions	Analysis and Problem Solving	0.659	1.02	0.10	0.630	1.20	0.17
Work that allows me to make lots of money	Personal Satisfaction	0.471	1.00		0.482	1.00	
Work that allows me to tell other people what to do	Personal Satisfaction	0.315	0.99	0.212	0.233	0.72	0.21
Work that allows me to have time with family	Personal Satisfaction	0.473	1.26	0.20	0.452	1.29	0.23
Work that makes people think highly of me	Personal Satisfaction	0.602	1.71	0.24	0.471	1.44	0.26
Work that is satisfying to me	Personal Satisfaction	0.580	1.30	0.19	0.557	1.40	0.24

consistency reliability of $\alpha = 0.61$. These alphas are lower than would be considered acceptable. However, given the small number of items on each of the subscales, low alphas are not unexpected. Implications for these findings are discussed below.

Conclusions

The present research found that the AWE Work Values scale assesses two separate sets of occupational values. One of these ("Analysis and Problem Solving") seems related to jobs in STEM fields while the other ("Personal Satisfaction") measures work values that could be met in any field, including but not limited to STEM careers. These findings should inform the way that those using the AWE Work Values scale analyze and interpret responses to the items. Our findings should also help to spur further investigation and development of scales that are specifically designed to measure occupational values related to STEM careers.

"Analysis and Problem Solving" Factor

Items on this factor examine students' valuing of thinking, problem solving, and use of STEM skills. Given the relationship between interest, efficacy, and valuing in occupational decision-making (Wang & Degol, 2013), this is an important consideration for those seeking to measure STEM occupational values. The ability to analyze and problem solve is consistent with the 21st Century Student Outcomes (Partnership for 21st Century Learning, 2011). These skills are also consistent with STEM professional training, such as the Accreditation Board for Engineering and Technology (2015) requirements for accreditation in applied science, computing, engineering, and engineering technology.

Although analysis and problem solving may not be unique to STEM careers, those who do not value these skills are unlikely to pursue a STEM career. If we view the items on the "Analysis and Problem Solving" factor from the perspective of expectancy-value theory (Wigfield & Eccles, 2000), we might view responses to these items as indicative of children's goals and schemata related to their future workplace. It is our position that ratings on this dimension are consistent with the developers' construct of interest in STEM careers. Despite this, because there are non-STEM careers involving analysis and problem solving, responses to these items would not likely be sufficient to predict interest in STEM careers. Future research should examine predictive relationships between valuing analysis and problem solving and career choice.

"Personal Satisfaction" Factor

The items on this factor are similar to the items included in many other occupational values tools, including the tool used by Duffy and Sedleacek (2007). In a study of college students over 10 years, Duffy and Sedleacek found that these types of values are unrelated to students having decided on a career. The values represented by the "Personal Satisfaction" factor might be consistent with many occupations, including STEM careers. Given the diversity of careers that fit under the umbrella of STEM, future researchers should examine the extent to which students believe various STEM careers are consistent with personal satisfaction. For students who do not value analysis and problem solving in their future careers, but who do value personal satisfaction, helping to make the connection between STEM and the ability to attain personal satisfaction may further influence interest in the field. Additionally, there may be other factors related to interest in STEM careers that are not included in the "Analysis and Problem Solving" or the "Personal Satisfaction" factors. Researchers developing tools intended to measure interest in STEM should consider the multiple values related to career paths in STEM. Focusing narrowly on one or two categories of work values may result in failure to identify other work values that have less obvious relationships to STEM careers.

Generalizability and Model Invariance

Although this research uses a large sample across multiple years and grades, the generalizability of our findings is somewhat limited. The data from the larger project were limited to a single geographic area, with a traditionally low-income sample comprised of primarily White and Black students. Future research on the factor structure of the AWE Work Values scale should examine the stability of our model in other contexts. Additionally, we did not examine model invariance across factors such as gender and ethnicity. Given the significant body of research on gender and ethnicity differences in interest in and attitudes about STEM and STEM careers (e.g., Dasgupta & Stout, 2014; Eccles, 2009; Snyder & Tan, 2006; Weisgram & Bigler, 2006), future research should examine model invariance for these and other potential factors.

Scale Reliability and Validity

One concern we have regarding both subscales is their low reliability. The use of short tools, such as Porfeli's (2007) or Sinsalo's (2004) work values scales, may result in occupational value dimensions that are measured with fewer than five items. Given the time and resources needed to implement a longer form, such as the commonly used Holland (1985) taxonomy, it is understandable that practitioners would be interested in using more abbreviated forms. As with the AWE Work Values scale, however, the use of only a small number of items on these scales results in poor reliability of measures. Schmidt (1996) acknowledges that this is a common problem for scales with few items. However, although low reliability is common for short scales, this does not alter the fact that a scale with low reliability will not measure a construct as accurately as one with strong reliability (Schmidt, 1996). Without sufficient reliability in responses, we cannot conclude that a scale is providing valid data. This proves problematic for those who want to measure occupational values with shorter scales with less than satisfactory Cronbach alphas. Given that the scales are likely to be used in the context of interventions in STEM areas, scales with lower reliability can present problems for statistical power. Heo et al. (2015) carried out a Monte Carlo study which found that, for a variety of designs (pretest–posttest, independent group comparisons), power was significantly reduced when Cronbach alpha and inter-item correlations were low. The lack of statistical power may lead to rejection of potentially effective STEM interventions for changing work values because there is not sufficient statistical power to find a difference.

In line with Schmidt's observation, we would recommend that future research focus on developing and testing additional items, particularly for the "Analysis and Problem Solving" factor. We continued to find a correlation between the "Analysis and Problem Solving" and "Personal Satisfaction" factors, indicating that, as currently measured, the two factors are not unrelated. By developing more distinctive items on the "Analysis and Problem Solving" scale, researchers may be better able to look at this as a separate and individual construct. Given improved and additional items, the measure could be more sensitive to changes in work values associated with STEM interventions.

Defining STEM Occupational Values

In addition to developing new items related to analysis and problem solving, researchers should continue to examine other values specifically related to interest and persistence in STEM careers. One area for further investigation is the role of task values as a function of career interest. When examining data from the Longitudinal Study of American Youth, Wang et al. (2015) found that math task value (e.g., "Math is useful in everyday problems") was a strong predictor of Grade 12 student interest in STEM careers.

Given the wide variety of careers that fall under the umbrella of STEM, researchers should also continue to examine occupational values that discriminate interest in different types of STEM careers. For example, in a meta-analysis of research articles related to career interest and gender, Su and Rounds (2015) found that women tended to be interested in STEM work environments with a people-orientation, while men tended to be interested in environments with a things-orientation. While this particular research examined differences based on gender, this may also be a factor for other cultural groups. Hidi et al. (2015) describe identity, especially gender and cultural identity, as an important factor in STEM career interest.

Finally, further research should examine the relationship between STEM occupational values and traditional occupational interest inventories such as the O*Net Interest Profiler (National Center for O*Net Development, n.d.). Foundational research in this area found a relationship between interests and values (Thorndike et al., 1968). However, more recent research has failed to find these relationships (Dobson et al., 2013). Dobson and colleagues suggest that the lack of evidence for a relationship between occupational values and interests might be a combination of greater precision in measuring interests (e.g., removing values items from interest scales) as well as the more informal approaches taken to measuring occupational values (e.g., card sorts, short scales). Without stronger ways of measuring occupational values, researchers are currently limited in their ability to identify relationships between values and interests. Additionally, it is unclear whether there might be a stronger relationship between STEM occupational interests and values more specific to STEM occupations, such as analysis and problem solving or math task values.

Summary

Although we believe the AWE Work Values scale developers intended the tool to measure STEM occupational values, we believe it actually measures two factors: personal satisfaction and analysis and problem solving. This information will help those using the AWE scale better interpret the data when determining whether program participants hold values consistent with STEM careers. While there are other tools that measure constructs consistent with the Personal Satisfaction dimension (Eccles, 1994), we have been unable to find tools that specifically examine to what extent students value the use of analysis and problem-solving skills in future occupations.

A reliable, valid tool that measures this set of occupational values would provide more detailed information for those researching and developing programs to influence STEM workforce participation. We believe the "Analysis and Problem Solving" factor identified in this research can serve as a foundation for just such a tool. If researchers and evaluators are interested in short scales that measure STEM career-related occupational values, we would suggest that researchers focus more on developing items such as those on the "Analysis and Problem Solving" factor.

Author Bios

Jessica M. Harlan is currently a Senior Research Consultant with Gallup Inc. Prior to that she worked as a Senior Program Evaluation Specialist at the Johns Hopkins University School of Medicine. She received a Ph.D. in instructional design and development from the University of South Alabama.

Email: Jessica_Harlan@Gallup.com

James P. Van Haneghan is a professor in the Instructional Design and Development Program at the University of South Alabama, where he teaches courses in educational research, educational measurement, and learning theory. The main focus of his research and evaluation work surrounds learning in the STEM disciplines.

References

Accreditation Board for Engineering and Technology. (2015). *Accreditation criteria*. Retrieved November 19, 2015, from http://www.abet.org/accreditation/accreditation-criteria/

Assessing Women and Men in Engineering. (2008). *Pre-college assessment resources.* http://aweonline.org/pre_college.html

Assessing Women and Men in Engineering. (n.d.a). *AWE FAQ page: Are all AWE surveys validated?* http://aweonline.org/faq.html#validate

Assessing Women and Men in Engineering. (n.d.b). *About us.* http://aweonline.org/default.html

Bandalos, D. L., & Finney, S. J. (2010). Exploratory and confirmatory factor analysis. In G. R. Hancock and R. O. Mueller, (Eds.). *Quantitative Methods in the Social and Behavioral Sciences: A Guide for Researchers and Reviewers*. New York, Routledge.

Brevik, K. M., Jean, K., Bowman, F. M., & Bowen, B. (2015). Impact of the You're Hired! program on student attitudes and understanding of engineering. Presented at the annual conference of the American Society for Engineering Education, Seattle, WA. Retrieved November 16, 2015, from https://www.asee.org/public/conferences/56/papers/13828/download

Byrne, B. M. (2010). *Structural equation modeling with AMOS* (2nd ed.). New York, NY: Routledge.

Castellan, N. J. (1992). Shuffling arrays: Appearances may be deceiving. *Behavior Research Methods, Instruments, & Computers*, *24*(1), 72–77.

Dasgupta, N., & Stout, J. G. (2014). Girls and women in science, technology, engineering, and mathematics: STEMing the tide and broadening participation in STEM careers. *Policy Insights from the Behavioral and Brain Sciences*, *1*(1), 21–29. https://doi.org/10.1177/2372732214549471

Deloitte Consulting LLP, Oracle, & The Manufacturing Institute. (2009). *People and profitability: A time for change*. Retrieved from http://www2.deloitte.com/content/dam/Deloitte/pt/Documents/manufacturing/pt(en)_dc_peopleprofitability_05112009.pdf

Demetry, C., & Sontgerath, S. (2013). Does a middle school intervention for girls have long-lasting differential effects on their perceptions of and participation in engineering? Presented at the annual conference of the American Society for Engineering Education, Atlanta, GA.

Dobson, L. K, Gardner, M. K, Metz, A. J, & Gore, P. A. (2013). The Relationship Between Interests and Values in Career Decision Making. *Journal of Career Assessment*, *22*(1), 113–122. https://doi.org/10.1177/1069072713492929

Duffy, R. D., & Sedleacek, W. E. (2007). What is most important to students' long-term career choices: Analyzing 10-year trend and group differences. *Journal of Career Development*, *34*(2), 149–163. https://doi.org/10.1177/0894845307307472

Eccles, J. S. (1994). Understanding women's educational and occupational choices: Applying the Eccles et al. model of achievement-related choices. *Psychology of Women Quarterly*, *18*, 585–609. https://doi.org/10.1111/j.1471-6402.1994.tb01049.x

Eccles, J. S. (2009). Who am I and what am I going to do with my life? Personal and collective identities as motivators of action. *Educational Psychologist*, *44*(2), 78–89. https://doi.org/10.1080/00461520902832368

Gonzalez, H. B. (2012). *An analysis of STEM education funding at the NSF: Trends and policy discussion*. Washington, DC: Congressional Research Service.

Harlan, J. M., Pruet, S., Van Haneghan, J. P., & Dean, M. D. (2014, June). Using curriculum-integrated modules to improve understanding of math and science content and STEM attitudes in middle grades students. Presented at the annual conference of the American Society for Engineering Education, Indianapolis, IN.

Heo, M., Kim, N., & Faith, M. S. (2015). Statistical power as a function of Cronbach alpha of instrument questionnaire items. BMC Medical Research *Methodology*, *15*, 86. https://doi.org/10.1186/s12874-015-0070-6

Hidi, S., Renninger, K. A., & Nieswandt, M. (2015). Emerging issues and themes in addressing interest in mathematics and science learning. Presented at the annual conference of the American Educational Research Association, Chicago, IL.

Holland, J. L. (1985). *Making vocational choices: A theory of vocational personalities and work environments* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Honey, M., Pearson, G., & Schweingruber, H. (Eds.)., (2014). *STEM integration in K–12 education: Status, prospects, and an agenda for research*. Washington, DC: The National Academies Press.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Maltese, A. V., & Tai, R. H. (2011). Pipeline persistence: Examining the association of educational experiences with earned degrees in STEM among U.S. students. *Science Education*, *95*(5), 877–907. https://doi.org/10.1002/sce.20441

National Center for O*Net Development. (n.d.). *O*Net Interest Profiler*. Retrieved November 20, 2015, from http://www.mynextmove.org/explore/ip

National Girls Collaborative Project. (n.d.). *Evaluation & Assessment*. Retrieved from http://ngcproject.org/evaluation-assessment

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instrumentation, and Computers*, *32*, 396–402.

Partnership for 21st Century Learning. (2011). *Framework for 21st century learning*. Washington, DC. Retrieved from http://www.p21.org/our-work/p21-framework

Pituch, K.A., & Stevens, J. P. (2016). *Applied multivariate statistics for the social sciences*. New York, NY: Routledge.

Porfeli, E. J. (2007). Work values system development during adolescence. *Vocational Behavior*, *70*, 42.60. https://doi.org/10.1016/j.jvb.2006.04.005

Rochester Institute of Technology. (n.d.). *Women in engineering: Improving program assessment tools*. Retrieved from http://www.rit.edu/affiliate/weimpact/links.php

Schmidt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, *8*, 350–353. https://doi.org/10.1037/1040-3590.8.4.350

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *Journal of Educational Research*, *99*(6), 323–338. https://doi.org/10.3200/JOER.99.6.323-338

Sinsalo, P. (2004). Changing work values and expressed educational plans of adolescents: A cross-sectional follow-up of three cohorts in Finland. *European Journal of Psychology of Education*, *16*, 227–236.

Snyder, T. D., & Tan, A. G. (2006). *Digest of education statistics 2005*. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

STEM Education: Industry and Philanthropic Initiatives: Hearing before the Subcommittee on Research of the Committee on Science, Space, and Technology of the House of Representatives. 113th Cong. 1 (2013).

Su, R., & Rounds, J. (2015). All STEM fields are not created equal: People and things interests explain gender disparities across STEM fields. *Frontiers in Psychology*, 6. https://doi.org/10.3389/fpsyg.2015.00189

Thorndike, R. M., Weiss, D. J., & Dawis, R. V. (1968). Canonical correlation of vocational interests and vocational needs. *Journal of Counseling Psychology*, *15*, 101–106.

Van Haneghan, J. P., Pruet, S., & Waltman, R. (2012, April). Issues in using a modified version of AWE's Middle School Survey in a study of an engineering design curriculum. Paper presented at the 2nd P–12 Engineering and Design Education Research Summit, Washington, DC.

Wang, M., & Degol, J. (2013). Motivational pathways to STEM career choices: Using expectancy-values perspective to understand individual and gender differences in STEM fields. *Developmental Review*, *33*, 304–340. https://doi.org/10.1016/j.dr.2013.08.001

Wang, M., Degol, J., & Ye, F. (2015). Math achievement is important, but task values are critical, too: Examining the intellectual and motivational factors leading to gender disparities in STEM careers. *Frontiers in Psychology*, 6. https://doi.org/10.3389/fpsyg.2015.00036

Weisgram, E. S., & Bigler, R. S. (2006). Girls and science careers: The role of altruistic values and attitudes about scientific tasks. *Journal of Applied Developmental Psychology*, *27*, 326–348. https://doi.org/10.1016/j.appdev.2006.04.004

Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, *25*, 68–81. https://doi.org/10.1006/ceps.1999.1015